The Challenge of Negation in Health Care Searches and Queries

Valerie J. Harvey, RT(R), PhD[1], Constance M. Ruzich, PhD[2], Jeanne M. Baugh, EdD[1], Bruce Johnston, MLS[3], Arthur J. Grant, PhD[2]
[1]Department of Computer and Information Systems, Robert Morris University, [2]Department of Communications, Robert Morris University, [3]RMU Libraries, Robert Morris University

Abstract: This poster deals with exclusionary queries implemented using the database language SQL and the VA FileMan database system and with retrieval searches involving negated concepts in medical narratives. The poster describes and presents error patterns in designing database queries, underlying comprehension issues regarding negative statements and queries, strategies and software for avoiding false positives in searches, and makes practical recommendations on identifying potential sources of error and avoiding incorrect or misleading results.

Description: Negation poses certain challenges for queries and searches used in health care. This poster deals with exclusionary queries implemented using the ISO database language SQL and the VA FileMan dialog-based interface and with retrieval searches involving negated concepts in medical narratives.

This project arose because instructors in undergraduate and graduate database courses noticed a large proportion of students making mistakes on certain queries and because library faculty noticed problems with searches involving negated concepts. The poster includes results gained from analyzing student performance on a four-query exercise designed for this project. Incorrectly designed queries can add incorrect rows to the result or omit correct rows or both, depending on the database and query. Professionals in various database production environments were asked for information on any problems encountered with exclusionary queries and on approaches used to avoid errors and this input is reported.

Relevant psycholinguistic and psychological findings are covered.[1] This poster explores underlying comprehension issues and makes practical recommendations on identifying potential sources of error and avoiding incorrect or misleading results.[2]

The portion of the project dealing with false-positive retrieval searches of medical narratives describes algorithms used to recognize negated concepts: *Negfinder* (Yale University School of Medicine)[3] and *NegEx* (University of Pittsburgh Center for Biomedical Informatics)[4]. The accurate identification of negated phrases is crucial to any automated retrieval system design in order to analyze these narrative reports correctly for clinical and research purposes.

Proposed actions include changes in general education, database training (for both SQL and VA FileMan), and orientation for conducting retrieval searches, strategies for review of database queries and retrieval searches used in clinical research, and advocacy of implementation of the new SQL:1999 standard, which offers improved syntax for formulating queries.[5]

References:

[1] Clark, HH. *Semantics and Comprehension*. Mouton, 1976.
[2] Wright, P, Bernard P. Effects of "More Than" and "Less Than" Decisions on the Use of Numerical Tables. Journal of Applied Psychology. 1975:60, 5: 606-611. Also Wright, P, Hull, AJ. Answering Questions about Negative Conditionals. Journal of Memory and Language. 1986:25: 691-709.
[3] Mutalik, PG, Deshpands; A, Nadkarni PM. Use of General-purpose Negation Detection to Augment Concept Indexing of Medical Documents: A Quantitative Study Using the UMLS. Proceedings of the American Medical Informatics Association 2001. 2001:598-609.
[4] Chapman, WW, Bridewell, W, Hanbury, P, Cooper, GF, Buchanan, BG. Evaluation of Negation Phrases in Narrative Clinical Reports. Proceedings of the American Medical Informatics Association Symposium 2001. 2001:105-109.
[5] Melton, J, Simon, AR. *SQL:1999: Understanding Relational Language Components.* Morgan Kaufman, 2002.